<u>IN THE SPECIFICATION</u>

Please amend the paragraph beginning at page 5, line 21, as follows:

IP AV fusion module 104 also makes a determination as to whether decoding of the spoken utterance will be accomplished using the merged acoustic-visual feature vectors $f_{A,V}(t)$ and acoustic-visual (AV) clusters or model (denoted in FIG. 8 1 as 108) or, due to a degraded visual environment, using only acoustic feature vectors $f_A(t)$ and acoustic-only (AU) clusters or model (also denoted in FIG. 8 1 as 108). This determination may be performed using a likelihood ratio test, to be described in detail below.